DOCUMENT RESUME

ED 299 572 CS 211 540

AUTHOR Low, Graham D.
TITLE Testing Reader-Writer Relations in Academic Text: Some Preliminary Considerations.
PUB DATE 87
NOTE 42p.; Paper presented at the Annual Meeting of the International Association of Teachers of English as a Foreign Language (22nd, Edinburgh, Scotland, April 11-14, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Audience Analysis; Communication Research; English for Academic Purposes; Higher Education; *Test Construction; Writing (Composition); *Writing Skills
IDENTIFIERS *Academic Discourse; *Author Reader Relationship

ABSTRACT

A. Criper and D. Davies (1986) have suggested that English for Academic Purposes (EAP) tests for prospective higher education students should include a component designed to measure the ability to manipulate reader-writer relations in academic texts. This paper examines the nature and scope of such manipulation and draws some preliminary conclusions about how it might be tested, both in terms of simplifying assumptions which could be made in specific situations and selecting test formats which appear particularly suitable. The process of manipulation is described from two points of view in the paper, firstly as negotiation of reader-writer power relations and assertions of identity and secondly, taxonomically, as a description of some of the strategies and devices available to the writer. The paper ends with a brief discussion of the results of version 1 of the University of York EAP test, part of which does try to measure the ability to control reader-writer relations.
(Twenty-two references and a sample test are appended.)
(Author/RAE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED299572

TESTING READER-WRITER RELATIONS IN ACADEMIC TEXT:

SOME PRELIMINARY CONSIDERATIONS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Graham Low
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Graham LOW
Language Teaching Centre
University of York
England YO1 5DD

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ABSTRACT

Criper and Davies (1986) suggested that EAP tests for prospective higher education students should include a component designed to measure the ability to manipulate reader-writer relations in academic texts. This paper is an attempt to examine the nature and scope of such manipulation and to draw some preliminary conclusions about how it might be tested, both in terms of simplifying assumptions which could be made in specific test situations and test formats which appear particularly suitable. The process of manipulation is described from two points of view, firstly as negotiation of reader-writer power relations and assertions of identity and secondly, taxonomically, by describing some of the strategies and devices available to the writer. The paper ends with a brief discussion of the results of version 1 of a University of York EAP test, part of which does try to measure the ability to control reader-writer relations.

CS211540

TESTING READER-WRITER RELATIONS IN ACADEMIC TEXT: SOME PRELIMINARY CONSIDERATIONS

Graham Low
Language Teaching Centre
University of York

1. INTRODUCTION[1]

One of the major aspects of writing any text is to gear it somehow to one's mental model of the intended reader(s). In their recent report on the British Council ELTS test, Criper and Davies (1986: 117-119) argue fairly forcefully that this is an important aspect of academic writing, which should be included in EAP tests for prospective higher education students and to which test designers should give rather more thought than they have done in the past. The purpose of this paper is to follow up Criper and Davies' suggestion and to examine, in a very preliminary way, what is involved in the concept of manipulating reader-writer relations in academic texts. The discussion will be followed by some comments based on our experience at the University of York of designing and piloting (regrettably on a very small group) a test which does try to test certain aspects of reader-writer

relations explicitly.

2. THE MANIPULATION OF READER-WRITER RELATIONS IN ACADEMIC TEXTS

2.1 General comments

People in academic life, whether students or staff, are faced with having to read and write a wide range of text types, from memos and notices, to the more specifically academic essays, conference papers, book reviews, textbooks, subject-specific newsletters and articles in a whole range of formal and informal academic journals. I shall take the view here that major overall differences in textual appearance and structure between these various types derive primarily from the nature of what the authors are trying to achieve in 'global' terms, but it is nevertheless true that a similarity of purpose and results has led to people creating category labels like 'textbook' or 'essay', to having certain expectations about what these will contain and even, in some cases, to codifying how exemplars should be produced (eg. the demand for specific types of referencing system). Given the range of texts which academics use and the corresponding importance of the written word in academic life, it is surprising how little published research exists concerning both the nature and the extent of methods used by academics to tailor their texts to the intended reader(s). There appears to be even less research on the ways in which genuine

readers (ie. not teachers of English composition) actually respond to writers' efforts in this area. In particular, there is a lack of ethnographic work of the sort done by Brown and Herndl (1986) with respect to English in a commercial environment. The discussion below owes much to previous work in ethnomethodology, to LePage's (1975) Act of Identity theory and to prototype theory (a brief overview of which can be found in Lakoff 1982). I have tried to avoid the rather complex terminology associated with linguistic work on prototype theory, but assume a basic familiarity with the other two sets of ideas.

2.2 Some text types and their writers and readers

In order to reduce this section to manageable proportions, only 3 text types will be considered here: the journal article, the textbook and the student essay. It is easiest to start with the formal journal article. The author is generally writing for other academics to read and comment. Although a distinction is generally made by writers in terms of whether they think they are introducing a new approach to the reader, or whether the reader is knowledgable in that particular area (Smith 1985), there is a general desire to see the reader as an equal; both participants are members of an academic, or subject-specific in-group. There may in fact be little basis for this assumption, but it is a myth that is strongly supported in practice. The relation between reader and writer is thus 'balanced', or 'symmetrical', and up-

down power relations are rarely allowed to intrude. The balance is generally extended to third parties whose work is referred to in the paper, and who must also in many cases be 'membershipped' as an equal. Not only must the membershipping not elevate or denigrate the third party, but just as importantly, it must not denigrate or unduly elevate either reader or writer. Indeed, failing to thus membership the reader or a third party can be used as a powerful tool for implicitly attacking them.

The writer of a textbook is in a rather different position. One only writes a textbook if one feels one has more knowledge than the reader and wishes to impart it. The knowledge relation between reader and writer, and hence the power relation between them and frequently the group membership relation, is therefore of necessity unbalanced or asymmetrical, with the writer 'up' and the reader 'down'. On the other hand, if this relation is made too overt, the reader will simply stop reading. The same will happen if the argument becomes too dense or the problem too difficult. Thus the writer really needs to create some sort of a bond between him/herself and the reader,though as the bond can rarely if ever lead to a fully balanced power relation, the result is frequently rather paternalistic.

The writer of an essay to be marked by a supervisor is in a different position again. The writer is a student and thus lower than the marker in the university career/power hierarchy and generally also as regards knowledge of the subject. The marker is an academic who is frequently well up in the subject and who has an according view of him/herself. To write an essay in the style

of a textbook would clearly reverse the power relationship and membership the marker as not belonging to the academic, or expert, community. This is not generally the best way to get a high mark! On the other hand, the writer of an essay is frequently made to feel that the ideal essay would be virtually publishable as a journal article, and that this is the goal to be reached. There is thus a clear tension in power terms. So, on the one hand, the writer is not an equal, accepted member of the academic hierarchy, and to write as though he or she was, could be taken as arrogance, insulting the reader, or even both! On the other hand, the expectation is that something resembling a journal article will in fact be produced, which implies that the writer must lay claim to being an in-group member. The difficulty of acceptably referring to research or opinions held by one's supervisor when writing a student essay can thus rapidly become acute. It should be noted that the writer is not alone in being pulled two ways. The marker frequently knows more about the topic than the student and has often worked in greater depth through the very problem which the student is trying to resolve. As a result, it is often hard to get interested in student essays, except where, and possibly unbeknown to the student, they make one rethink one's own position. Against this, the marker must create an illusion, by pretending not to be an expert in the field concerned and imagining how convinced he/she would be in such a case by the essay being marked. So, although the last thing one might actually need in real life was explicit details of all the steps leading to conclusion X, one would nevertheless

penalise a student for lack of clarity if they were in fact omitted. Moreover, the addition of role tension on the part of the reader, to a situation already made difficult by role tension on the part of the writer, can turn the question of discussing and evaluating the marker's own research into something which both parties can at times experience great difficulty in resolving satisfactorily.

The need to manipulate reader-writer relations extends to all aspects of text construction. In a student essay, for example, the writer must decide, amongst other things,

1. which facts/ideas are relevant, given a known marker,
2. how explicit the argument should be,
3. which studies to mention,
4. what the permissable limits of praise or objection are, given the marker, and what precise degrees of either ought to be given to those studies and their findings, selected in this particular instance,
5. how much (and what sort) of background to give,
6. whether full references should be given,

This brief selection makes it quite clear that the interpersonal aspect of academic writing can determine aspects of 'content' and 'organisation'. It is not a question of a dichotomous break between style and content, with reader-writer relations functioning simply to provide markers designed to humanise otherwise 'objective' text (as Roe 1977 appears to suggest). It

is also clear that selection of the quantity and quality of evaluative language is in part determined by interpersonal considerations (as Peters 1986 has pointed out). Lastly, it is equally clear that interpersonal aspects of writing include reference to third parties and not just to reader and writer. In this connection, Smith (1986) distinguishes usefully between extra-text relationships (eg. Student-Professor) and temporary text-internal relationships (eg. questioner-responder). As long as the analysis allows for students and professors to have unclear relations on several levels, with the student unable to predict totally which will come into play at which part of the essay (ie. role tension is accepted), Smith's point allows a sensible generalisation. This is that the average student is likely to have only text-internal relations with third parties, whereas the expert marker, journal article and textbook writer will also have text-external relations. Moreover, the student writer has to try and write with this in mind.

2.4 Devices available to the writer

It is perhaps useful to envisage the writer as having three types of strategy or device available for manipulating interpersonal relations. Selecting ideas, or choosing which to footnote (v. Decisions 1-3) may be considered as 'organisational' decisions. Strategies such as repeating particular terms are perhaps better seen as 'rhetorical' ones. For example, Brandt (1986) found the

same student writer making more use of the teacher's terminology and avoiding elaboration in a class assignment, but increasing the use of explanations and evaluations, and using synonyms rather than repetitions, when it was a question of writing an Arts essay for a wider audience who shared less knowledge of the situation leading up to the writing of the text. Again in the context of Arts-type essays, Peters (1986) found that there was a stronger correlation between ~~use of~~ control over evaluative language and good grades, than control over personal reference and grades. In addition to organisational and rhetorical strategies, writers also have access to a wide variety of overt lexicogrammatical features of English, eight of which are outlined below.

(a) Terms of address

Table 1 lists a number of exponents for addressing reader and writer in academic texts. They can be arranged in two surprisingly parallel lists and both sets are more or less scalable in terms of the discourse situation postulated by the writer. At the top is a 'here and now' deictic situation, involving definite live people who can comment and be addressed in a direct fashion. By the middle of the scale, the restriction on definite specified participants has gone, the participants are downgraded to third parties and there is increasing distance from the deictic situation. By the bottom of the scale, there is still greater distance and neither definite nor live persons at all.[2]

PRESENCE OF READER	PRESENCE OF WRITER
Remember when?	I, We (= I)
You (personal)	We (you and me) We (him/her and me)
We (you and me)	
One	One
You (general)	You (general)
We (general)	We (general)
The reader	The writer
	The present author (=myself)

Table 1. Some exponents used to mark personal reference.

It is obvious that the terms in the two lists can also be used to manipulate both asserted and implied relations between reader and writer, though the exact nature of the effects obtainable is far less clear.

On admittedly very little hard evidence, I would suggest that 5 degrees of distancing might be isolated, in the sense that they would be interpreted fairly consistently across a range of text and reader types. The 'scale' runs from[(1)] explicitly asserting

a relationship (as when the writer talks individually to the reader as a person, e.g. "you") through (2) implying a relation (when the writer avoids describing the reader directly as "you", but allows the use of inclusive "we"), to (3,4) admitting that the person exists, but avoiding a specific relationship (when the writer either (3) places the reader into a universe of likely readers and calls him/her "one", or else (4) refers to his/her function alone (eg. "the reader", without stressing that an actual person undlerlies the reference)). The end of the scale is virtual total distancing, (5) where no explicit reference is made to either reader or writer (eg. "It ought to be obvious at this point that" = I think you should by now realise that.....).

This scale, or if it is preferred, these combinations, permit the use of three apparent absurdities. Firstly, the absence of reference to either reader or writer is tantamount to the pretence either that neither exists, or else that the text exists meaningfully independent of them (Lakoff's 1982 'objectivist myth'). Secondly, labelling onself as 'the writer' is to act as if one was totally separate from, and independent of, oneself; a situation it can be hard to maintain for long. Thirdly, the use of 'we' where no plurality is logically possible, seems justifiable only by reference to the royal 'we', with the writer elevating him/herself over the reader to godlike proportions.

Above and beyond the occurrence of possible absurdities, the lists and the relationship scale raise two interesting points. The first is the conclusion that references to reader and writer

and references to third parties are closely linked. The argument here shows one merging into the other in terms of deictic distance. THe argument in Section 2.2 suggested a link in terms of power relations and insults. It remains to point out that these two links, deictic distance and power relations, are themselves linkable. Roe (1977: 36) cites the following, from a sixth-form textbook:

> "A student sometimes thinks that all the information required to be known about a reaction is contained in the simple, balanced equation for the reaction, but this is a very limited and false outlook."
>
> (Brown 1964: 306)

The writer is attempting to insult a class of readers to which the present reader might belong, but yet not to directly identify the particular reader with that group. The use of third-party reference permits the use of group-inclusion techniques, while maintaining deictic distance and the absence of direct specific reference.

The second interesting point is the question of whether any terms in Table 1 are actually neutral with respect to reader-writer relations. The obvious candidates are terms like 'I' and 'the writer'. But even here, use of 'the writer' may well be used to signal a positive unwillingness to elevate oneself above the reader, while the use of 'I' can indicate the opposite. A clear-cut example would be a preference for 'I' rather than 'we' by the writer of a textbook. Nevertheless, Smith's (1985) point that 'you' generally marks greater interactiveness than 'I' (for the

simple reason that 'You' also implies the existence of 'I') seems fair enough.

(b) Modal and auxiliary verbs

The writer has a range of possible options, many of which are likely to be interpreted as referring to reader-writer power relations. Several such verbs can be ordered roughly in terms of what we may call 'default-value' coercion (since the precise degree of coercion felt will depend upon the verb following, and various aspects of the context):

must
will have to
should
has to / will have to
may like to
may

Several verbs ideally require the reader to think carefully about the precise meaning intended by the author. For example,

> "In Part IV we shall turn our attention to the sociology of culture".
>
> (Bottomore, 1962: 129)

If this was written in a student essay, rather than in a textbook, the reader might well feel that the idea of futurity was less salient than that of coercion. Similarly, expressions like "We would expect" or more problematically,

> "We should expect changes in caste to be greater in urban areas than in the villages".
>
> (Bottomore, 1962: 186)

leave it uncertain what precise balance between coercion, logical necessity, and implications about the reader's intellectual abilities is intended.

At this point in the argument, it is sufficient to note firstly, that modals can have several concurrent meanings (ie. the situation is rarely one of 'either this meaning or that' (Lakoff 1977 and others)) and secondly, that the reader's interpretation of particular verb phrases employing modals is likely to vary depending on the text-type involved. That is to say, specific modal verbs do not have an inherent, genre-free reader-writer relation associated with them.

(c) Group inclusion devices

I have labelled these in functional, rather than pseudo-logical or syntactic terms, as the argument is getting complex enough by this point. There are a number of terms in English which allow

the writer to categorise the reader as belonging to the group of people who think X or Y. These are words like 'most' or 'all'. The degree of certainty with which the reader is included is also open to infinite variation (eg. "Many people, of which the reader is perhaps one, frequently admit to feeling ...." versus "We have all at times felt .....").

The three sets of lexicogrammatical devices so far discussed do not occur in isolation; they tend to co-occur, as in the examples cited above, in complex verbal groups. This very complexity, however, leads to a number of difficulties when it comes to measuring how far learners are able to control them, a point which will be taken up again in the following sections.

(d) Personal epithets

The range of possible personal epithets is enormous, but their use is closely geared to how far the writer sees the reader and him/herself as being members of the same group. Referring to someone as "Ojemann (1984)", involving the absence of all descriptive tags, tends to be used where Ojemann is considered as an equal member of the academic community, along with the reader. It is thus the norm in journal articles. The use of "Professor G. Ojemann, writing in ..." might well be read as marking the fact that the reader is not being considered as a full member of the academic in-group, where titles are rarely if ever used between members, but is not so far down the power hierarchy as to need

greater explanation. On the other hand, expressions such as "the well-known neurosurgeon George Ojemann is on record as saying" suggest that the reader is being considered as totally outside the academic community and way down the power hierarchy and hence needs to be told that Ojemann is in fact well-known (at least in the writer's possibly biassed view). A reasonably clear distinction may perhaps be made between descriptive/eulogistic/derogatory labels and name/title labels. As anecdotal support for this, one may cite the fact that postgraduate students in particular frequently spend much of their university life worried about how to refer in different contexts to the friendly, but world famous professor and pro-vice chancellor that colleagues and secretaries alike call 'John Smith', or just 'John'. They rarely worry about whether to add 'the eminent Austrian economist' on the front.

(e) Evaluative language

Disjunct adverbials like 'surprisingly' are very often vague with respect to whether it was the writer who was surprised or the reader who should be. Roe (1977) and Smith (1986) both add that authorial comments on the relevance or validity of an argument, such that X is "important", or topic Y has been "almost unconsidered", also have inter-personal effects, presumably since the author is informing the reader that a judgement is being made on his/her behalf. Smith's claim that since these terms occur in all scientific texts, relative frequency may be used as an indicator of textual interactiveness, seems too strong as a

general hypothesis, since a reader may willingly permit the decision in, say, a basic primer, and perceive little interaction, but react strongly and negatively to the same thing in a technical paper or student essay. Moreover, all such terms are susceptible to considerable downtoning modification, which would add to relative frequency, but may serve to decrease interactiveness.

(f) Interrogatives and imperatives

Smith (1986) distinguishes usefully between less interactive rhetorical questions and more interactive 'genuine information questions', pointing out that the latter tend to be used where the audience is known (eg. a committee working paper). He also distinguishes three types of imperatives: (1) demands to third parties, (2) imperatives demanding a state of mind or minimal action ("Let us assume", or "Assume for a moment") and (3) more interactive commands requiring definite action (eg. "Move the paper"). This still leaves the question of conditions unaccounted for, and without research into reader reactions it is hard to know just how interactive readers perceive statements like "If it is accepted that ....." to be.

(g) Deixis

Genuine dexis is not likely to used much in academic writing except in scripts of oral presentations or memos to colleagues. However, as Smith (1986) again points out, metaphoric dexis, which he labels 'text time' and 'text space', which postulate an immediate and direct interaction between reader, text and writer, are common. An example of text time would be "We may pause now and ....", with a text-space version being "We may pause at this point in the argument and ...". Terms like 'the above' are also examples of text space.

(h) Passive mood

Hopper and Thompson (1985) consider that in English, as in many other languages, the primary function of full verbs is to offer a report of an activity that someone or something performed. Passives clearly permit selective focus by highlighting the activity and downgrading the actor or agent responsible. Thus passives tend to lower interactiveness. However, one cannot conclude that this is their only function. Tarone et al. (1981), studying two astrophysics journal articles, noted that passives were contrasted with actives to mark downgrading not so much of the agent, but of the activity itself. Moreover, there were two degrees of downgrading involved. Firstly, it was used as a third party group exclusion device, to mark less interesting work by other researchers against which the authors were reacting. Secondly, the same authors used it as an irrealis marker, when

referring to work that had not yet been done, but would, or might, be done in the future. Interestingly, this is not the end of the story, since there is evidence for passives being used as group-inclusion devices via a somewhat different rationale. Brown and Herndl (1986), looking at report writers in a commercial corporation, found writers with less job security overusing passives and complex nominalisations, despite knowing that both led to reader distance and 'poor style', as an attempt to identify themselves with the better writers, who they felt had greater value, authority or job security within the company. The writers concerned simply saw both linguistic devices as salient attributes of the academic style used by the group they wished to identify with; in this case, the importance of passives lay in the fact and frequency of their use, rather than the nature of what they referred to.

2.5 The role of controversy

The various devices and their combinations are frequently more suited to certain tasks or contexts than to others. To take an example:

> "This (partial differential) equation is of the form known as separable, so we can write (10) as
>
> $$\frac{\partial P\ (s,t)}{P\ (s,t)} = \lambda(s-1)\ \partial t$$

and we get

$$\ln P(s,t) = \lambda(s-1) + c$$

$$P(s,t) = e^{\lambda t(s-1)}\, e^{c}\ "$$

(A Statistics textbook, ref. lost)

The reader is unlikely to feel any objection about being included in the group of people who would perform such an operation and agree with such a transformation. The complete derivation is not provided, but the point is that it could be. Although the writer could have used "Writing (10) as.... gives ...", the ability to introduce a personal element can be exploited to positive effect, thus rendering slightly more human(e?) an otherwise highly abstract task. On the other hand, if the following were to appear in a textbook or an essay,

> "We may begin by considering the numerous dichotomous classifications which have already been mentioned on several occasions"
>
> (Bottomore 1962: 113)

> "We may disregard here the difficulties of functionalist explanation..."
>
> (Bottomore 1962: 195)

> "We may conclude at this point by noting the fact that

the caste system is very closely connected to specific features of the Hindu religion"

(Bottomore 1962: 185 adapted)

the situation is very different. Examples have been chosen which are not likely to be read as particularly coercive in this context, though the degree of perceived coercion may well still differ considerably from reader to reader. The point at issue here is the question of group membership. Few readers are likely to object to "We may begin", if only from the realisation that one must start somewhere and the precise starting point selected is often unimportant. The chances of the expert reader agreeing that functionalist explanations are better omitted, or that Point Y is the place that they would themselves stop the argument, however, are much lower. It is thus quite interesting to note that 'We may begin' is not an exact counterpart to 'We may conclude'. The crucial differentiating factor would seem to be the degree to which the reader is likely to disagree with the proposition he or she is being associated with.

A similar set of arguments can be given concerning the use of 'I' in texts. Using 'I' clearly demarcates the author from the rest of the in-group as well as from members of the out-group(s). This might well be appropriate in most contexts where an author wishes to show that a particular view is a personal opinion and goes against received wisdom. It might also be accepted where the author considers him/herself to be pre-eminent in the field and the reader accepts this, as when a famous professor of economics

writes, concerning recent developments in China,

> "These (trends) have occurred as I predicted they would 5 years ago..... I have noted that this trend is sufficiently clear and will not be reversed"
>
> (Cheung 1986)

We may conclude this section (sic) by noting that (1) particular expressions are suited to certain local (as well as global) tasks or contexts, (2) that a text type like an essay or a textbook is likely contain a variety of such local tasks, and (3) there is thus no reason to expect that the occurrence or omission of particular expressions, such as 'I', is a clear function of text type. The selection of appropriate exponents by the writer remains educated inductive guesswork based on at least three things, considered concurrently: knowledge of the text-type, knowledge of what the text is trying to achieve at a particular point, and the mental model of the reader. The writer may thus choose to change from more personal to less personal styles within a text, however formal that text may be felt to be.

The argument has so far focussed on the importance of reader-oriented style changing, but thinking about the reader also leads to presssure for imposing limits to change and a need for some consistency. On the question of local consistency, there is a need for some research into the scope of the term 'local' in different situations. I suspect that different readers have different opinions, but as a rule of thumb (which could easily be

tested) I suggest that most readers would prefer to be labelled consistently, whether by 'you', 'we', or 'one', within an orthographic sentence.

3. GENERAL PROBLEMS IN TESTING READER-WRITER RELATIONS

3.1 The outline presented so far shows that the art of manipulating reader-writer relationships is a complex one which (a) can be considered on a number of levels, (b) can involve a wide range of devices and strategies, not all of which involve overt surface markers and (c) involves a notable absence of clearly bounded, non-overlapping categories. Any comprehensive test of the ability to control reader-writer relations in academic texts would seem require a situation involving a model of one or more readers, a writer with a purpose or intention towards the reader(s), one or more text-types, a set of subtasks for which the writer will have 'local' rather than just global intentions towards the reader, and what I have elsewhere described as a developing context (Low 1982).

3.2 The fact that reader-writer relations can affect virtually all aspects of a text, means that it can prove hard to isolate manipulating them as a skill, or set of skills, to be tested. In practice, a number of simplifying assumptions are going to have to be made. Editing, rather than composition from zero, may help

to provide a solution here, since tests involving editing allow a coherent context to be provided while, at the same time, permitting points of interest to be isolated reasonably effectively in several different ways.

3.3 The role of the reader is an area likely to cause particular difficulties in a test of manipulating reader-writer relations. I have tried to show that coping with role tension is an important aspect of the skill, and that tension between roles affects both reader and writer. Indeed, it may well be true that the problem is most acute in student essays, which is likely to be one of the main foci of such tests. It manifests itself in various ways. Firstly, many essays are written with one particular reader in mind and this can cause difficulties where multiple and anonymous markers are employed. Secondly, the use of teachers of English as markers can introduce a seriously inauthentic element since, if they do not themselves supervise postgraduates, or conduct research, they are unlikely to experience the sort of role tension being discussed.

3.4 The finding by Smith (1986: 115) that a number of published academic papers establish in the opening sections the limits of the informality to be used in the paper as a whole, suggests that tests involving journal articles at least could profitably make use of opening sections. Against this is the difficulty mentioned in Low (1982) that testing stylistic aspects of language would seem to require some sort of context to have been previously

established. This however, suggests that there should be a textual lead-in before the question, which may well mean testing stylistic aspects of reader-writer relations at the end, rather than at the start, of a test.

3.5 Despite the great complexity of the topic, it is possible for a writer to adopt very simple solutions which can produce acceptable texts in many situations. Thus the student who uses the maxim 'Be completely impersonal in all essays and never mention either reader or writer' may produce apparently very good essays, but nevertheless be very poor at manipulating the full range of devices and strategies available. A test which simply asked that student to 'write an essay' would provide little indication of precisely where the student's real problems lay. A comprehensive test would seem to need several different texts and/or for the testee to have to modify them to match differing sets of specific criteria.

3.6 If a number of simplifying assumptions are made, it does seem, as Smith (1985) claims, that frequency counts can provide rough quantitative estimates of the relative interactiveness of texts, or parts of texts. Three such assumptions are (1) that boosting/downtoning does not affect interactiveness significantly, (2) that it is not necessary to conduct the analysis in terms of the negotiation of power relations and (3) that modal verbs will have isolatable and roughly similar interactiveness values from text to text. To the extent that

these assumptions do not detract greatly from what one wishes to test in any particular situation, the ability to rank order texts with respect to interactiveness in a principled way is likely to be extremely useful to someone trying to construct a multi-text test design.

3.7 The finding that writers who are insecure members of a company, say, persist in using 'poor style' as an attempt to demonstrate group membership, could prove a problem for language testers. Part of the difficulty might be removable by thinking carefully about the details of the intended reader which are given to the candidate, but in general the topic seems unresearched and no clear solutions are available.

3.8 It has been mentioned several times that most actual tests of reader-writer relations are unlikely to be comprehensive ones examining control across all relevant text types. The focus is frequently going to be on whether a prospective student can write an essay. It should be clear, however, from the discussion in Sec. 2 that, although this does permit the use of all sorts of simplifying assumptions, it does not mean that essays can be divorced totally from other text-types. A more realistic and fruitful approach would be to test how far the candidate can avoid writing like the textbooks which probably form the major component of his/her reading and the major input to the essay concerned.

4. SOME EMPIRICAL COMMENTS: THE R.E.S.T.

4.1 The Rapid English Screening Test was produced by the Language Teaching Centre of the University of York in 1986 to establish whether students coming onto its summer pre-sessional EAP courses needed particular help with academic writing at intermediate level. It consists of 5 subtests. The first 4 are used for placement purposes; the last one, which contains more difficult reading material as input, has so far been used primarily for diagnostic purposes. This fifth question has two sections; one (Q5a) testing control over metaphors of argument, the other (Q5b)the ability to manipulate reader-writer relations. The design of the complete test is described in Low (1986b). Although the number of testees is still small, (n=27, of whom 18 completed Q5b), the 1986 results have thrown up a number of points of general interest. The test was administered as students arrived for the course and again as they finished it. The following comments relate to the post-course results, which were marked by myself, as testees had by this time some familiarity with the test format and the scores reflect their abilities at the point where they began their academic courses.

4.2 The format of Q4, Q5a and Q5b is that of writing a student essay. The intended reader/marker is explicitly profiled in various parts of the test as a Professor with expert knowledge of the field and as someone sensitive to the way in which he is 'talked to' in essays. The passage chosen for testing reader-

writer relations was the Introduction to the essay, with Q5a providing in effect a textual lead-in to Q5b. Q4 involves manipulating a data-based section from the middle of the essay. For two reasons, it was decided to sequence the passages so that the test ends with the Introduction. Firstly, it was felt that a context needed to be built up before stylistic items were tested (Low 1982; Lee and Low 1982). Secondly, writers do frequently edit or write introductory sections last, and this way a more natural reflection of the writing process could be achieved. Although considerable care was taken to show that Q5b concerns the Introduction (v. Appendix A), 3 out of 18 testees rewrote it as a Conclusion, even though it was the second time they had taken the test. The effect of several explicit statements and examples was apparently over-ridden by one single use of the word "final" at the point where the testees had to start writing ("Write your improved version of the final paragraph here").

The implication would seem to be that sequences of subunits of text, in genres like an academic essay, can create very strong expectations in some students, and these expectations can override test instructions in cases where the tester chooses to break the anticipated text sequence. The fact that the three students concerned had fairly low scores on all five test questions may indicate that linguistically less proficient students rely particularly heavily on this sort of textual expectation.

4.3 The R.E.S.T. has to be adinistered within 90 mins to 2 hours and marked very rapidly. It also has to contain an initial

section establishing whether certain fairly basic aspects of English grammar are known. In order to achieve this, a number of simplifying assumptions were made. Firstly, no organisational aspects of reader-writer relations were included. Secondly, only certain phrases were considered (by the test designer) as having inter-personal values. It was thought that one of the by-products of restricting the topic to the stylistic manipulation of chunks interpolated into a text would be that the focus of the test task would become more clearly defined. Testing chunks in this way requires that the text in which the chunks are embedded be given, so testees were only asked to edit a pre-existing Introduction. However, they were asked to actually rewrite it and they were not told which chunks were problematic. What happened in practice is instructive. Many testees appeared not to appreciate the validity of the argument itself. This seems to have sprung more from a desire to remove certain statements, rather than merely to rephrase the passage in their own words, though the two may have been connected. Whatever the reason, many test scores were contaminated by varying degrees of summarisation, which proved impossible to remove completely. There are two conclusions here. Either the chunks need to be identified for the testees, or the testees should be asked, not to rewrite the paragraph, but to mark their corrections on top of the existing text. The latter solution might also remove some of the temptation to turn an Introduction into a Conclusion.

4.4 The task in Q5b was in effect to find and improve the 8

problematic chunks of text. Six chunks concern references to the reader or the writer, one involves references to a third party and one is an evaluation of part of the argument in Q4. The reference to the third party involves evaluatory epithets as well as titles, and is considered here as constituting two separate, if juxtaposed, items. The six first and second person references involve different degrees of verbal complexity, but were all of necessity complex (subject plus verb group) items. Table 2 contains the relevant details.

TABLE 2 HERE

The results are of interest in two ways. Firstly, only one testee even attempted to modify the evaluatory Item 4. It may well be that the emphasis in the instructions on 'the tone' of the paragraph and the suggestion that the Professor 'won't like being talked to like this' was responsible. Possibly it was the absence of an overt marker such as 'in my view' which led to Item 4 being overlooked. A third explanation might be that testees normally expect test items to be totally self-contained. Item 4 raises a real problem for testing, however. Personal references can be manipulated without extensive discussion of extra-text situations. Evaluatory language, however, frequently requires either detailed knowledge of the subject area or, as here, knowledge of a specific pre-existing argument. Either way, it is difficult to construct items which direct the testee's attention sufficiently clearly, which are answerable by ~~all but~~ other than the best-

ITEM	SENTENCE	PRONOUN	MODAL VERB	GROUP INCLUSION DEVICE	EPITHET	FACILITY VALUE
1	1	We	must (agree)	-	-	.58
2	1	We	will	all (be concerned)		.53
3	3	You	will (need to examine)		-	.64
4	3	(A Q4a argument is erroneously labelled 'conclusive')				0
5a	3	-	-	-	the eminent Californian psychologist	.28
5b	3	-	-	-	Dr.E. (Rosch)	.17
6	4	We	would	all (agree)	-	.47
7	4	You	- (start)	-	-	.47
8	5	You	should (define)	-	-	.53

n=18 α = .87

Table 2. Summary of Q5b items

read testees and which are not extremely time-consuming.

The second point of interest is to examine how the testees responded to the personal reference items. From the facility values and an examination of the test papers, I conclude rather tentatively that:

(a) Modals like 'must' and 'will' were perceived as clearly infringing the maxim that one must not order the reader around.

(b) Modals like 'will' and 'would', with their greater range of possible values, were found harder to interpret; testees found it harder to see when interpersonal effects were likely to be inappropriate or counterproductive.

(c) The group inclusion term 'all' was seen as a clear infringement of the maxim that one should not actively force the reader to belong to the same group as the writer. Of 36 possible occurrences in the 18 test papers, the word was only retained on 3 occasions; in some cases removing 'all' was the only change made.

(d) Many testees failed to even notice that E. Rosch was (glaringly) inappropriately described, while others did try to modify the terms, but failed. Only 3 managed to resolve both items successfully. It looks as though two things happened. Firstly, several students just did not consider

that reference to third parties was an important feature of the text. Secondly, of those who did, most had great difficulty in predicting the likely effects of different types of wording. Quite possibly they did not see third party references as being covered by the same sort of ground rules as references to the reader and the writer.

4.5 A particular problem for the establishment of rating scales for individual test items derives from the comments made earlier about the role of local, rather than global, consistency in reader-writer relations. This is the independence of possible responses to different items. Items 1 and 2, which are both in the same sentence, provide a good illustration of the problem:

"We must agree that in this essay we will all be very much concerned with definitions...."

One possible, if not ideal, resolution for Item 1 would be "It seems to me that (this essay must necessarily be concerned with definitions...)". One quite acceptable resolution for Item 2, given that this is a student essay, would be "....I will be concerned in particular with the question of definitions". If both solutions are applied together, however, the sentence becomes much less acceptable, suggesting uncertainty and incompetence on the part of the writer: "It seems to me that I will in this essay be concerned in particular with the question of definitions". It is hard to see how problems like this could

be resolved on anything but an ad hoc basis.

4.6 The last point of interest also concerns rating scales; it is the question of how far answers to particular items are likely to be scalable, and if they are, how many degrees of acceptability should one (or do people in practice) recognise? The responses to Item 8 illustrate the difficulty of applying a three-point rating scale (Acceptable, Clearly Unacceptable, and Unacceptable but in the Right Direction). Item 8 was "You should define love operationally". Under 'Acceptable' were classed,

"Love needs to be defined operationally"
"Love is best defined operationally"
and "It is preferable to define love operationally",

while under 'Clearly Unacceptable' were included,

"You must define love operationally"
and "You will have to define love operationally".

This leaves a large number of expressions somewhere in the middle. My own reaction as a marker is that

"Love can be defined operationally",
or "I define love operationally",

while not perfect, are nevertheless more acceptable than

"I would define love operationally"

or "Love could be defined operationally".

The most one can conclude at this stage is that more research is needed.

5. CONCLUSION

Even a brief examination of the nature of reader-writer relations shows that this is a highly complex area which can pervade almost all aspects of writing. Smith (1985, 1986) succeeded in providing quantitative measures of text interactiveness based on a componential type of analysis. However, the introduction of the reader as a person and the translation of the writer's purpose into a task-by-task attempt to manipulate power relations make it hard to maintain the clearly-bounded non-overlapping categories needed for what Luke (1986) would call 'a black box model of communication'. While Smith wished primarily to account for the distribution of overt lexicogrammatical features in given texts, the language tester is faced with serious problems by needing in many cases to go beyond this. The present paper has suggested a number of ways in which simplifying assumptions can sometimes be made by a tester at the design stage and has broached several areas which would repay more research. Three areas of particular interest for test design are: (1) controlling reference to extra-test

(or extra-text) knowledge, (2) examining how real academic markers resolve problems of role tension and (3) observing how testees react to various types of multi-text test designs (some ideas for the latter can be found in Low 1986a and in press). In the absence of this sort of research, Criper and Davies' suggestion remains, as Cyril Weir has aptly put it, an intriguing, but intractable, "can of worms".

G.D. Low
20 May 1987

NOTES

1. An earlier version of this paper was presented at the Colloquium on Language Testing Research: Priorities and Prospects, University of Reading 6-9 April 1987. The present version has benefited considerably from comments by the participants, in particular by Cyril Weir and Bernard Spolsky.

2. Strictly speaking, the top of the scale is better analysed as a prototypal situation involving a collection of attributes. Downgrading can take place along any of the dimensions represented by the attributes. A linear scale is a reasonably accurate representation in this case, however, since discourse prominence, deictic distance and 'full humanity' seem to be particularly closely linked.

REFERENCES

Bottomore, T.B. 1962. Sociology. London: Unwin.

Brandt, D. 1986. 'Text and context: how writers come to mean'. In B. Couture (ed).

Brown, G.I. 1964. Introduction to Physical Chemistry. London: Longman.

Brown, R.L. and Herndl, C.G. 1986. 'An ethnographic study of corporate writing: job status as reflected in written text'. In B. Couture (ed), pp 11-28.

Cheung, S. 1986. 'Is China going capitalist?' South China Morning Post

Couture, B. (ed). 1986. Functional Approaches to Writing. London: Pinter.

Criper, A. and Davies, D. 1986. Edinburgh ELTS Validation Project: Final Report. London: British Council.

Hopper, P. and Thompson, S. 1984. 'The discourse basis of lexical categories in universal grammar'. Language 60(4): 703-752.

Lakoff, G. 1977. 'Linguistic Gestalts'. Proceedings of the Regional Meeting of the Chicago Linguistic Society. Chicago: CLS.

" 1982. Categories and Cognitive Models. Working Papers, Series A, No.96. Trier: L.A.U.T.

Lee, Y.P. and Low, G.D. 1982. 'Towards a classification of tests of language use: the dimension of test directness.' Journal of Applied Language Study 1(1): 16-28.

Le Page. 1975/1977. 'Sociolinguistics and the problem of

'competence'. In V. Kinsella (ed.) Language Teaching and Linguistics: Surveys. Cambridge: CUP, pp 39-59.

Low, G.D. 1982. 'The direct testing of academic writing in a second language'. System 10(3): 247-257.

" 1986a. 'Storylines and other developing contexts in use-of-language test design'. Indian Journal of Applied Linguistics 12(1-2): 15-38.

" 1986b. 'Beyond the REST: a brief introduction to the York Rapid English Screening Test. Language Testing Update 3: 5-13.

" In press. 'Appropriate design: the internal structure of course units.' In The Second Language Curriculum, edited by R.K. Johnson. Cambridge: Cambridge University Press.

Luke, K.K. 1986. 'Dialogue between dialogues: the management of dual participation frameworks in the telling of stories in conversations in Cantonese'. Presented at the 'Erving Goffman: an Interdisciplinary Appreciation' conference. University of York, 8-11 July.

Peters, P. 1986. 'A study of dominant function in academic writing'. In B. Couture (ed), pp. 175-185.

Roe, P. 1977. Scientific Text. ELR Monographs. Birmingham: University of Birmingham.

Tarone, E., Dwyer,S., Gillette, S. and Icke, V. 1981. 'On the use of the passive in two astrophysics journals'. ESP Journal 1(2): 123-139.

Smith, E.L. 1985. 'Functional types of scientific prose'. In Systemic Perspectives on Discourse, edited by J.D. Benson and W.S. Greaves. New York: Academic Press, pp 241-257.

" 1986. 'Achieving impact through interpersonal component'. In B. Couture (ed.) pp 108-119.

TEST 5

Situation: The main part of the essay is finished and you are writing the Introduction.
You show it to an English student for comments.
Here are her overall comments:

> Not too bad. I have marked a couple of spelling errors + a few problems with metaphor. The tone of the last paragraph is wrong. It is not hard to correct, but I suggest you do correct it.

Your job: Paragraphs 1 and 2: Add in more appropriate expressions, where the English student suggests.

Paragraph 3: Rewrite it as suggested above.

WRITE YOUR ANSWER IN THIS COLUMN

THE PSYCHOLOGICAL BASIS OF ROMANTIC LOVE IN HUMANS.
IS THERE ONE?

Sec.1. Introduction

Although it is fairly clear that there is some sort of phenomenon to study, as millions of pounds are spent annually in Europe and the USA on love stories

Prof. will not like being talked to like this

We must agree that in this essay we will all be very much concerned with definitions, in particular whether certain features are criterial and serve to define romantic love wherever it is found. Sarsby (1983) interprets Stone's (1977) 5 factors as such a list. You will need to examine several studies (eg. the conclusive findings of Richards 1983) to see whether there is empirical support for this, or for the alternative, possibly more appealing, position taken by the eminent Californian psychologist, Dr. E. Rosch, which would predict a series of factors likely to be associated with love, none of them, however, being strictly necessary. We would thus all agree that it makes little sense defining love before you start the enquiry. You should define love 'operationally', by using studies which ask people what they mean by love, or by analysing commercially successful romantic stories.

WRITE YOUR IMPROVED VERSION OF THE FINAL PARAGRAPH HERE

.....

.....

.....

.....

1988 ADDITIONAL NOTE

This paper describes Version 1 of the R.E.S.T. and its description in Language Testing Update. I am at present analysing the results of Version 2, which involves editing, rather than extended writing (at least in the Reader-Writer relations subtest). The sampling of the items for Version 2 was taken from pp 1-26 of this paper and thus represents a broader selection than we used for Version 1. The stimulus text has also been completely rewritten and represents a considerable improvement over the one included here. There is a new introductory booklet describing Version 2 of the whole test. We are at present working on Version 3 and trying to devise some tests in order to arrive at a 'natural' marking scheme for some of the items.

GDL 15 Mar. 1988